Burning money and (pseudo) first-mover advantages

Huck, S.; Müller, W.

Published in:
Games and Economic Behavior

Publication date:
2005

Document Version
Publisher's PDF, also known as Version of record

Available online at www.sciencedirect.com
SCIENCE DIRECT
Games and Economic Behavior 51 (2005) 109–127
GAMES and Economic Behavior
www.elsevier.com/locate/geb

Burning money and (pseudo) first-mover advantages: an experimental study on forward induction

Steffen Huck [a], Wieland Müller [b]

[a] *Department of Economics and ELSE, University College London, Gower Street, London WC1E 6BT, UK*
[b] *Department of Economics, Tilburg University, Postbus 90153, 5000 LE Tilburg, The Netherlands*

Received 5 August 2002

Available online 25 August 2004

Abstract

The mere potential for one player to burn money prior to play has been shown in theory to be an effective device to select this player's most preferred outcome, e.g., in the battle-of-the-sexes game [J. Econ. Theory 48 (1989) 476, J. Econ. Theory 57 (1992) 36]. In this study we assess the behavioral relevance of this theoretical claim. It is shown that its validity depends on whether the game is played in extensive or normal form. In extensive form games first movers benefit substantially from having the opportunity to burn money in advance. The effect is much weaker in the normal form game. However, in two control treatments first movers who can select between two related games gain an advantage although standard forward induction arguments do not have any bite. These results suggest that we need to make further theoretical advances to understand the role of physical timing and first-mover advantages in games.

JEL classification: C72; C92

Keywords: Forward induction; Burning money; Battle-of-the-sexes game; Coordination; Experiments; Physical timing; First-mover advantage

E-mail addresses: s.huck@ucl.ac.uk (S. Huck), w.mueller@uvt.nl (W. Müller).

0899-8256/$ – see front matter
doi:10.1016/j.geb.2004.03.006

1. Introduction

As innocent as the "battle-of-the-sexes" (BoS) game shown in Table 1 may seem, it poses a challenge to game theorists as there is no obvious prediction about how agents should play this game. After all, the game has three different Nash equilibria—two asymmetric ones in pure strategies, (P, Y) and (Y, P), and one in mixed strategies in which each player plays Y with probability $1/4$ and earns an expected payoff of $3/4$. To overcome this indeterminacy without entering into the realm of equilibrium selection theory, game theorists have suggested to look at slight variations of the basic game—variations that are supposedly more realistic and lead to a unique solution.[1]

- A first approach, that has been introduced by Farrell (1987), is to allow players to engage in nonbinding pre-play communication. Here actual play of the BoS game is preceded by one or more rounds in which players simultaneously make announcements regarding their intended play. Farrell shows that this game has a symmetric equilibrium in which players coordinate more often than in the mixed equilibrium of the simple BoS game.
- A second approach involves giving one player, say the row player, an outside option that gives her a certain payoff in case she decides not to play the game. Now suppose the row player chooses to play the game nevertheless. Does this convey any information that is carried forward into the ensuing subgame? Some authors have argued it does (Kohlberg and Mertens, 1986). Because why should the row player enter into the BoS game if she were to expect a payoff lower than her outside option payoff? Thus, the fact that the row player refuses to take the outside option tells the column player something about the row player's intention in the BoS game to be played next, namely that she wants to earn a higher payoff than she would have earned by exercising the outside option. (Otherwise she could simply have taken it.) This principle is called *forward induction.*[2] It signifies the idea that the intended play of an agent at the beginning of a subgame might be inferred from her play in earlier stages of the game. In the game at hand, an outside option payoff greater than 1 and smaller than 3 for the row player would be sufficient to ensure coordination on her preferred equilibrium (P, Y).
- A third approach (van Damme, 1989, and Ben-Porath and Dekel, 1992), on which we will focus in this study, involves, surprisingly, to give one player the opportunity

Table 1
Battle-of-the-sexes game

		Column player	
		$Y(ield)$	$P(referred)$
Row	$Y(ield)$	0, 0	1, 3
player	$P(referred)$	3, 1	0, 0

[1] Equilibrium selection theory à la Harsanyi and Selten (1988) would select the mixed equilibrium.

[2] Note that this principle corresponds to the notion of *strategic stability* in normal form games introduced by Kohlberg and Mertens (1986).

Table 2
The extended game in which the row player may burn one payoff unit in the frist stage

Game to be played after the row player's decision not to burn money

		Column player	
		Y	P
Row player	Y	0, 0	1, 3
	P	3, 1	0, 0

Game to be played after the row player's decision to burn money

		Column player	
		Y	P
Row player	Y	−1, 0	0, 3
	P	2, 1	−1, 0

to "burn" some money before the game is played. Suppose that the row player may or may not burn 1 "util" before the game begins. Then the actual game that is played is the one shown in Table 2 where the right subgame is derived from the BoS game by substracting one unit from the row player's payoff in each of the four cells. Again one can apply some forward induction reasoning though it is this time a bit trickier: Suppose the row player decides to burn money. Then she must expect that (P, Y) will be played because the two other equilibria give her less than what she could have secured for herself in the original game.[3] Hence, if this is correctly understood by the column player, the row player can, by burning money, secure herself a payoff of 2. In the next step of reasoning, this implies that if she decides not to burn money she also must expect that (P, Y) will be played (because both other equilibria give her less than what she could have secured for herself by burning money). Hence, the subgame perfect equilibrium based on forward induction is for the row player not to burn money followed by choosing P in both subgames and for the column player to choose Y in both subgames.

All three models are designed to put the original game in a context in which it becomes easier for players to coordinate. So one question is whether real people benefit from these contexts as much as fully rational players do. In other words, one might ask whether these models help to predict actual behavior. Laboratory experiments seem ideally suited to answer this questions. The available evidence on the success of forward induction is mixed. With respect to the three theoretical approaches to resolving the BoS game the literature provides us with the following findings.[4]

[3] In the original game the row player's expected payoff cannot drop below $3/4$.

[4] Further experimental evidence on forward induction can be found in Brandts and Holt (1992) who report data from several two-stage games with complete information. Their findings support the usefulness of forward induction. Cooper et al. (1992) investigate a 2×2 coordination game with two Pareto-ranked equilibria and report that a payoff-relevant outside option changes play in the direction predicted by forward induction. Van Huyck et al. (1993) report the success of forward induction in a setup in which the right to participate in a coordination game is auctioned off prior to play. Brandts and Holt (1995) report—in the context of a BoS game with an outside option for one of the players—that forward induction works well where it coincides with a simple dominance argument, but fails where dominance does not apply. Cachon and Camerer (1996) investigate a setup in which subjects may pay a fee to participate in a coordination game with Pareto-ranked equilibria. They report that play is consistent with forward induction. Finally, Balkenborg (1998) reports that in a simple game in extensive form in 80% of all cases play resulted in the outcome predicted by backward induction whereas the outcome predicted by forward induction was only rarely observed.

- Cooper et al. (1989) provide experimental evidence on the role of pre-play communication in the BoS game. They investigate three different communication structures: one-way communication with one round of messages and two-way communication with either one round or three rounds of messages concerning intended play. They conclude that "[c]ommunication significantly increased the frequency of equilibrium play" and that "[o]ne-way communication was most effective in resolving the coordination problem" (p. 568). In particular, the frequency of ex post equilibrium play rises from 48% in the basic BoS game to 95% in the BoS game with one-way one-round communication.[5]
- Four years on, Cooper et al. (1993) also provide experimental evidence on the role of forward induction in BoS games with outside options. They report that "[t]hough the presence of the outside option changes play, [they] find only limited support for the forward-induction hypothesis" (p. 1303). However, the frequency of equilibrium play (conditional on reaching the BoS subgame) rises from 41 to 90% whereas the frequency of the row player's preferred equilibrium (again, conditional on reaching the BoS subgame) rises from 19 to 90%.
- Regarding the third approach mentioned above, the experimental literature has been silent up to now. In this study we want to complement the experimental evidence by implementing the possibility of money burning.

To obtain a benchmark, we ran sessions on the standard BoS game as shown in Table 1. To assess how the potential of burning money affects behavior, we next ran sessions on the extended game shown in Table 2. To learn more about the forces driving behavior, we also conducted sessions on the normal form representation of this game (see below, Table 3). Note that deleting iteratively weakly dominated strategies in the normal form game leads to the same prediction as the forward induction argument in the sequential game. Finally, we conduct two additional control treatments with a game tree identical to the extended game from Table 2 but payoffs chosen such that forward induction does not select an equilibrium.

Our results concerning the simple BoS game replicate those reported earlier (e.g. by Cooper et al., 1989): Play is reasonably well predicted by the symmetric mixed equilibrium. Regarding the behavior in the sequential money burning game we find that the option to burn money significantly changes play in the BoS subgame—to roughly the same extent as Cooper et al. (1989) report for an outside option: The equilibrium preferred by the row player (who has the opportunity to burn money) emerges in 69% of all cases in which the BoS subgame is reached. In contrast to these results, we find that theory has almost no predictive power in the normal form representation of the money burning game: The predicted outcome is played in only 6.5% of all cases. The reason for this seems to be that subjects do not iteratively eliminate dominated strategies but stop after one round of reasoning.[6] This shows that the actual timing in the game is important for whether or not the opportunity to burn money conveys an advantage.

[5] For the effects of pre-play communication on behavior in coordination games see, e.g., Cooper et al. (1992), Clark et al. (2000), Charness (2000), and Burton et al. (1999).

[6] A stunning failure of subjects to go through longer chains of reasonings is reported in a recent paper by Kübler and Weizsäcker (in press) on informational cascades. For further evidence on subjects' depth of reasoning

The results of our additional control treatments highlight the role of sequential play in a surprising manner. There the first mover also gains a significant advantage although forward induction does not predict such an advantage. This suggests the necessity for making theoretical advances in understanding the role of physical timing and first-mover advantages.

The remainder of the paper is organized as follows. Section 2 presents the experimental design. The results of the experiments are reported in Section 3. Section 4 introduces two control treatments and discusses the literature on the role of physical timing in games. Section 5 concludes.

2. Experimental design and procedures

The experiments were computerized[7] and conducted at the University of London in November 2000, July 2001, and March 2003. Subjects were students from various fields. Upon arrival in the lab, participants were assigned to a computer terminal and received written instructions. After reading them, questions could be asked which were answered privately.

We employ a design that is quite similar to the one used in Cooper et al. (1993).[8] In all, we ran five treatments. For each treatment we conducted five sessions and in each session a group of 10 subjects interacted repeatedly. Thus, altogether 250 subjects participated in the experiment. Each session consisted of 30 rounds which was common knowledge among participants. Also, subjects knew that they would be randomly matched in pairs in each of the 30 rounds. The numbers given in the payoff tables were measured in a fictitious currency unit called "Points". The fixed exchange rate of £1 for 4 Points was commonly known. In addition to their earnings, subjects received a £5 show-up fee.

We proceed by describing the three treatments we initially set out with. After we had analyzed the data of these treatments, we decided to add two further control treatments. Mimicking this process, we shall introduce the fourth and the fifth treatment in a separate section further below.

In treatment BoS subjects played the standard BoS game as shown in Table 1. In this treatment no labels were assigned to participants. The instructions simply used the words "you" and the "other participant". In each round all subjects had to decide between two options ("Option 1" and "Option 2"). After one round's play, subjects were informed about the choice of the participant they were matched with and about their individual payoff as well as their total payoff so far.

see the seminal work of Nagel (1995) or the more recent papers by Costa-Gomes et al. (2001) and Weizsäcker (2003), as well as the literature cited therein.

[7] We used the software tool kit *z-Tree*, developed by Fischbacher (1999).

[8] Cooper et al. (1993) report results of five treatments: (i) BoS, the baseline treatment with the standard battle-of-the-sexes game; (ii) BoS-300, introducing an outside option that (according to forward induction) is high enough to select the first mover's preferred outcome; (iii) BoS-300-N, the same as (ii) in normal form; (iv) BoS-100, the same as (ii) but with an outside option too low to select the first mover's preferred outcome; (v) BoS-SEQ, the standard BoS game where the row player moves first in physical time but her choice is not observed by the column player.

In treatment *Burn* subjects played the sequential money burning game where we labelled the two roles by "A" and "B". Prior to each odd round subjects were randomly assigned to one of the roles. In even rounds the roles were reversed. Matching was random in each round and these procedures were commonly known (for details, see the instructions in Appendix A). At the beginning of a round, an A had to decide whether to go "to the right" or "to the left". By doing so, A selected a table. One table corresponded to the payoff table of the standard BoS game, the other to the payoff table that results from burning money. Furthermore, subjects were instructed that B would be informed about A's decision regarding the two tables and that afterwards A and B had to decide simultaneously about their options in the relevant table. After each round, subjects were informed about the choice of the participant they were matched with in the respective subgame. As in treatment BoS, subjects were also informed about their individual payoff and their total payoff so far.

In treatment *Norm* subjects played the (reduced) normal form of the money burning game. Role assignment (as a row or a column player) and matching was as above. Subjects decided simultaneously in their respective roles. Again, after completion of a round, subjects were informed about the choice of the participant they were matched with and about their individual payoff as well as their total payoff so far. The corresponding 4-by-4 payoff table is shown in Table 3. Here, the first component of the row player's strategy indicates whether she burns money (*Burn*) or not (*Not*) and the second component indicates which action she takes afterwards. Similarly, the first entry in the column player's strategy indicates this player's choice after the row player did not burn money and the second entry refers to the choice after the row player did burn one payoff unit. (In the experiment the actions were simply labelled Row and Column 1, 2, 3, and 4.) The solution of this game can be obtained by iterative elimination of weakly dominated strategies: $((Not, P), (Y, Y))$ is the only surviving strategy profile after 5 steps.[9]

Due to possible losses in treatments Burn and Norm, subjects in all three treatments got an initial endowment of 20 Points (the show-up fee mentioned above). The monetary payoff was determined by individual earnings in the 30 rounds (plus the initial endowment).

Table 3
The normal-form representation of the money burning game

		Column player			
		Y, Y	Y, P	P, Y	P, P
Row player	Not, Y	0, 0	0, 0	1, 3	1, 3
	Not, P	3, 1	3, 1	0, 0	0, 0
	$Burn, Y$	−1, 0	0, 3	−1, 0	0, 3
	$Burn, P$	2, 1	−1, 0	2, 1	−1, 0

9 *Step 1:* $(Burn, Y)$ is weakly dominated by (Not, P) for the row player. *Step 2:* (Y, P) is weakly dominated by (Y, Y) and (P, P) is weakly dominated by (P, Y) for the column player. *Step 3:* (Not, Y) is strictly dominated by $(Burn, P)$ for the row player. *Step 4:* (P, Y) is weakly dominated by (Y, Y) for the column player. *Step 5:* $(Burn, P)$ is strictly dominated by (Not, P) for the row player. As a result, the strategy vector $((Not, P), (Y, Y))$ is the only one surviving the deletion of iteratively dominated strategy.

Whereas sessions in treatments BoS and Norm lasted about 30 minutes, a session in treatment Burn lasted approximately 45 minutes. Average earnings in treatment BoS, Burn and Norm were £12. 40, £ 14. 52 and £12 58, respectively.

Summarizing the theoretical predictions, we formulate the following hypotheses:

Forward Induction Hypothesis. There will be a significant first-mover advantage in treatment Burn. Row players will typically choose not to burn money and (P, Y) will be the most frequently observed outcome.

Iterative Dominance Hypothesis. There will be a significant first-mover advantage in treatment Norm. The most frequently observed outcome will be $((Not, P), (Y, Y))$.

3. Experimental results

In the following we focus on experienced and settled-down behavior from the second half, i.e., the last 16 rounds of the experiment.[10] We take an even number of rounds since this gives us for each of the subjects an equal number of decisions in each of the two player roles. Furthermore, for all subsequent statistical tests we treat each session as one independent observation.

Let us first consider treatment BoS. As mentioned above there is no reason to expect subjects to coordinate on one of the asymmetric equilibria. In fact, as shown in Table 4, strategy P was chosen in 65% of all cases which is not too far away from the mixed-equilibrium frequency of 75%. Table 4 also shows the probability distribution over the four cells of the payoff matrix as implied by the observed behavior.[11] Each of the two asymmetric equilibria are expected with a probability of 22.75%.

Cooper et al. (1993) who rely on an identical BoS game[12] report that strategy P was chosen in 63.5% of all cases. Thus, our results are remarkably close to theirs.

Now consider the results of treatment Burn in which the row player has the opportunity to burn one payoff unit prior to the play of the simultaneous-move game. Recall that theory

Table 4
Treatment BoS: Relative frequency of choices and outcomes in the last 16 periods

$N = 400$		Column player		Total
		Y	P	
Row	Y	12.25	22.75	35.00
player	P	22.75	42.25	65.00

[10] While there are some significant time trends in the first half of the experiment, there are no significant correlations between time and behavior in the last 16 rounds. Moreover, all our results regarding differences between treatments are robust against selecting different subsets of periods.

[11] Note that the actual distribution would give a slightly distorted picture as it depends on how participants were matched with each other.

[12] In their experiments, however, all numbers were multiplied by 200.

Table 5
Treatment Burn: Absolute and relative (in parentheses) frequencies of outcomes in the last 16 periods
Left (right) panel after the row player's decision not to burn (to burn) money

$N = 375$		Column player		Total
(93.8)		Y	P	
Row player	Y	30 (8.0)	11 (2.9)	40 (10.9)
	P	258 (68.8)	76 (20.3)	334 (89.1)
	Total	288 (76.8)	87 (23.2)	375 (100)

$N = 25$		Column player		Total
(6.2)		Y	P	
Row player	Y	2 (8.0)	3 (12.0)	5 (20.0)
	P	37 (52.0)	7 (28.0)	20 (80.0)
	Total	15 (60.0)	10 (40.0)	25 (100)

predicts the row player to choose *not* to burn and subsequent coordination on the row player's preferred equilibrium. Inspecting Table 5, which summarizes the data, we observe that in 93.8% of all cases row players indeed choose not to burn money. Next, we observe that the incidence of the row player's preferred equilibrium rises from 22.75% in treatment BoS to 68.8% in treatment Burn. This increase is statistically significant at $p = 0.005$ (one-tailed Mann–Whitney U-test based on five observations per treatment). Nevertheless, this is still a smaller change than predicted by theory. Row players now play P in 89.1% of all cases (compared to 65% in BoS) and column players' relative frequency of Y increases from 35 to 76.8%.

Let us next compare earnings in treatments BoS and Burn. Subjects in treatment BoS earn on average 0.92 Points (standard deviation based on session means 0.12). In contrast, in treatment Burn row players earn on average 2.00 Points (std. dev. 0.54). According to a one-tailed Mann–Whitney U-test this is a significant increase vis-à-vis subjects' earnings in treatment BoS ($p = 0.004$). Column players' average earnings, however, are at 0.78 (std. dev. 0.07) significantly lower than average earnings in the baseline treatment BoS ($p = 0.028$). Finally, a one-tailed Wilcoxon matched-pairs test reveals that row players' earnings in treatment Burn are significantly higher than column players' earnings ($p = 0.0215$).[13]

We summarize what we found so far by stating

Result 1. There is support for the Forward Induction Hypothesis. Players significantly benefit from having the option to burn money.

Let us now turn to treatment Norm. Recall that according to the Iterative Dominance Hypothesis, we expect behavior to converge to $((Not, P), (Y, Y))$. But inspection of Table 6, which shows data from the normal form treatment, reveals that the hypothesis is dramatically falsified: Play coincides with the predicted outcome in only 6.5% of all cases.

Regarding the behavior of the row player we notice that—taken together—the (implicit) decision not to burn money accounts for 94.5% of all cases which is close to the 93.8%

[13] Again, the test is based on session averages; i.e., each session serves as one observation for both, row's and column's earnings.

Table 6
Treatment Norm: Absolute and relative (in parentheses) frequencies of outcomes in the last 16 periods

$N = 160$		Column player				Total
		Y, Y	Y, P	P, Y	P, P	
Row player	Not, Y	4 (1.0)	21 (5.3)	5 (1.3)	21 (5.3)	51 (12.8)
	Not, P	26 (6.5)	170 (42.5)	40 (10.0)	91 (22.8)	327 (81.8)
	$Burn, Y$	– (–)	2 (0.5)	1 (0.3)	1 (0.3)	4 (1.0)
	$Burn, P$	3 (0.8)	8 (2.0)	2 (0.5)	5 (1.3)	18 (4.5)
	Total	33 (8.3)	201 (50.3)	48 (12.0)	118 (29.5)	400 (100)

in treatment Burn. Furthermore, strategy (Not, P) which is the row player's part of the predicted outcome was chosen in 81.8% of all cases (which compares to 83.6% in Burn[14]).

As subjects fail to go through the whole process of iterative elimination of weakly dominated strategies, it seems worthwhile to check how far they actually got. They seem to master the first step: The strategy that is deleted in the first round, $(Burn, Y)$ of the row player, is almost never chosen (1%). After the elimination of $(Burn, Y)$, the column player's strategies (Y, P) and (P, P) become dominated. So, if subjects manage to reason two steps, they should be played rarely. However, they are chosen in 79.8% of all cases. This is a clear failure of rationality, especially as subjects constantly switch roles and know both sides of the game.

Note, however, that the row player's strategy $(Burn, P)$ that makes (Y, P) for the column player worse than (Y, Y) and (P, P) worse than (P, Y) is also very rarely chosen, namely in only 4.5% of all cases.[15] Cooper et al. (1993) observed a similar failure of the iterative-deletion argument. In their treatment as well as in ours, the iterative deletion of dominated strategies stops after the first round.[16]

Recall that subjects in treatment BoS earned on average 0.92 Points. In treatment Norm, however, row players earn on average 1.52 Points (std. dev. 0.60) which appears to be due to the high frequency (42.5%) of the outcome $((Not, P), (Y, P))$ which gives the same payoffs as the strategy vector that survives the iterative deletion of weakly dominated strategies. Column players in treatment Norm earn on average 0.72 Points (std. dev. 0.07). While the increase in row players' earnings is only significant at $p = 0.076$ (one-tailed

[14] Note that in treatment Burn the row player first chooses the subgame and then an action in this subgame whereas in treatment Norm these two choices are made simultaneously. In order to make the two cases comparable, in treatment Burn we multiply the frequency with which row players chose P in the BoS subgame, 89.1%, with the relative frequency with which the BoS subgame was selected, 93.8%. Hence the number $83.6\% = (89.1 \times 93.8)/100\%$.

[15] This is what evolutionary game theory would suggest: Weakly dominated strategies can survive if the strategies against which they are doing badly disappear. See, for example, Nachbar (1990) and Samuelson (1993).

[16] Similar findings are also reported in other studies (see footnote 6).

Table 7
Treatment Norm, Alternative representation: Absolute and relative (in parentheses) frequencies of outcomes in the last 16 periods. Left (right) panel after the row player's decision not to burn (to burn) money

$N = 378$		Column player		Total
(94.5)		Y	P	
Row player	Y	25 (6.6)	26 (6.9)	51 (13.5)
	P	196 (51.9)	131 (34.7)	327 (86.5)
	Total	221 (58.5)	157 (41.5)	378 (100)

$N = 22$		Column player		Total
(5.5)		Y	P	
Row player	Y	1 (4.5)	3 (13.6)	4 (18.2)
	P	5 (22.7)	13 (59.1)	18 (81.8)
	Total	6 (27.3)	16 (72.7)	22 (100)

Mann–Whitney U-test), the decrease in column players' earnings is highly significant ($p = 0.004$).

For the sake of comparison with the results in treatment Burn, Table 7 shows an alternative representation of the results in treatment Norm. Here the results are presented as if subjects had played the sequential game. Comparing the results of the left subgame in Tables 5 and 7, it is apparent that the row players' behavior is very similar whereas column players yield less often in treatment Norm. In any case, we can summarize by stating

Result 2. There is no support for the Iterative Dominance Hypothesis. However, row players weakly benefit from having the option to burn money.

What we found so far is that there is some support for the Forward Induction Hypothesis as the option to burn money drives behavior in treatment Burn into the direction predicted by this theory. However, iterated deletion of dominated strategies fails dramatically in the normal form game. These differences concerning play in the extensive and the normal form representation suggests that the fact of giving only one of the players the opportunity to burn money in treatment Burn might have created a focal asymmetry favoring the row player. To assess the explanatory power of this hypothesis, we conducted two control treatments in which forward induction is not an issue. These are described in the following section.

4. Two control treatments

4.1. Design and predictions

The games used in our two control treatments are shown in Tables 8 and 9. In both games forward induction has no bite.[17]

[17] This is obvious in the case of the second game shown in Table 9. Here the first choice is materially irrelevant since both subgames are identical. To see that forward induction has no bite in the first game in Table 8, observe that the payoff in the mixed strategy equilibrium is identical in both subgames.

Table 8
The game in treatment Control I

Game to be played after the row player's decision to go "Left"

		Column player	
		Y	*P*
Row	*Y*	0, 0	1, 3
player	*P*	3, 1	0, 0

Game to be played after the row player's decision to go "Right"

		Column player	
		Y	*P*
Row	*Y*	0, 0	1.2, 2
player	*P*	2, 1.2	0, 0

Table 9
The game in treatment Control II

Game to be played after the row player's decision to go "Left"

		Column player	
		Y	*P*
Row	*Y*	0, 0	1, 3
player	*P*	3, 1	0, 0

Game to be played after the row player's decision to go "Right"

		Column player	
		Y	*P*
Row	*Y*	0, 0	3, 1
player	*P*	1, 3	0, 0

The games are designed to assess the relevance of two behavioral explanations for why first movers may have an advantage in the main treatment Burn. The first explanation (which has been suggested by a referee) is a variant of the standard forward induction argument which we shall call *behavioral forward induction*. The idea is that subjects do use forward induction but ignore mixed strategies.[18] Consider the game in our first control treatment, shown in Table 8 and assume that subjects only consider payoffs that correspond to pure-strategy equilibria. Then subjects in the control treatment may reason that by choosing left a row player could either hope to earn 1 or 3, whereas by choosing right she could earn 1.2 or 2. Therefore, a subject acting as a row player and choosing left might signal that she will go for the payoff of 3 since the alternative equilibrium gives a payoff that is lower than both equilibrium payoffs in the right subgame. Obviously, this version of behavioral forward induction also works in the original game of treatment Burn.[19] However, it does not work in the second control treatment where the row player's initial choice is basically irrelevant since both subgames are materially identical (only rows and columns have been switched).

While behavioral forward induction predicts a first-mover advantage in the first but not the second control treatment, the alternative behavioral explanation predicts first-mover advantages in both settings. This explanation is based on the idea that the fact that one

[18] Recall that the forward induction argument in the money burning game of Table 2 rests on the fact that the row player's expected payoff in the BoS subgame cannot drop below 3/4. This payoff is equal to the row player's maximin payoff and the row player's payoff in the mixed equilibrium.

[19] One can argue that "behavioral forward induction" works better in the control treatment than in treatment Burn where it requires subjects to think one level deeper. In Burn subjects have to figure out that by not burning money the row player signals that she wants a payoff of at least 2, the payoff that she could obtain by burning money.

player can make a first choice renders this player's preferences "focal" (even if the choice is materially irrelevant). This alternative focal-point explanation is related to the effects of what Rapoport and collaborators called "positional order protocol".[20] Under a positional order protocol players choose sequentially in physical time although information about previous choices is not revealed. Game theoretically such sequences are equivalent to simultaneous decision making.[21] Behaviorally, however, there is evidence for a substantial and significant advantage in moving first even if information about choices is not passed on. In fact, Cooper et al. (1993) were the first to conduct such a positional-order-protocol experiment with the two-player BoS game and the preferred equilibrium of the player who moved first in physical time emerged in 62% of all cases.[22] Similar results are presented in Muller and Sadanand (2001). A first-mover advantage in our second control treatment would prove an even stronger impact of physical sequences in playing games. After all, when subjects reach one of the two subgames perfect symmetry between the players is restored. This is different under a positional-order-protocol where the game theoretical symmetry is broken by physical time. In our experiment it would be an event *in the past* that breaks symmetry in the present. Thus, a first-mover advantage in our second control treatment would stress the importance of incorporating physical time into a game theoretic framework.

We summarize the predictions for the control treatments in the following three divergent hypotheses:

Forward Induction Hypothesis 2. There will be no first-mover advantage in both control treatments.

Behavioral Forward Induction Hypothesis. There will be a significant first-mover advantage in the first but not the second control treatment.

Physical Timing Hypothesis. There will be a significant first-mover advantage in both control treatments.

---

[20] See Rapoport (1997) and Güth et al. (1998), as well as the references cited in both.

[21] von Neumann and Morgenstern (1947) distinguish between "anteriority" (chronological order of play) and "preliminarity" (priority in information). They note that both coincide if and only if the extensive form game has perfect information. A more formal approach to this idea is due to Thompson (1952) who formulated transformations that preserve the reduced normal form of an extensive game, e.g., the interchange of moves principle.

[22] We decided not to run such a treatment for two reasons: First, as mentioned in footnote 12, the BoS game used by Cooper et al. (1993) corresponds to ours except that their payoffs are ours multiplied by 200. Thus, there is no reason to expect that we would observe different behavior and, therefore, add any new insight. Second, we believe that a treatment better suited to shed additional light on the behavior of subjects in our experiment, is one in which we keep the basic structure of the money-burning game, i.e., the row player has a "first choice" prior to the *simultaneous* play of a subgame. Note again that in the treatment in Cooper et al. (1993) using the positional order protocol, players in the BoS (subgame) move *sequentially*.

4.2. Results

We conducted five sessions for each of the two control treatments. In each session again a group of 10 subjects interacted repeatedly. Apart from the different payoffs in the right subgame, everything else was exactly the same as in treatment Burn. The results of the first control treatment are presented in Table 10, the results of the second in Table 11.

Inspection of Table 10 showing the results of treatment *Control I* reveals the following facts. First, the original BoS subgame is chosen much more often than the alternative right subgame (in 89% of all cases which compares to 93.8% in treatment Burn). Second, the incidence of the row player's preferred equilibrium rises from 22.75% in treatment BoS to 82.9% in the control treatment which is highly significant.[23] It appears that this effect is even more pronounced than the one observed in treatment Burn where the row player gets his preferred equilibrium only in 68.8% of all instances of the BoS subgame but it turns out that the difference between Burn and the first control treatment is not statistically significant.

Next, let us examine Table 11 that shows the results of treatment *Control II*. Here we do not expect that one subgame is chose more often than the other (after all they are virtually identical). The data are close to fifty–fifty with the left subgame being chosen in roughly

Table 10
Treatment Control I: Absolute and relative (in parentheses) frequencies of outcomes in the last 16 periods

$N = 356$		Column player		Total
(89.0)		Y	P	
Row player	Y	12 (3.4)	2 (0.6)	14 (3.9)
	P	295 (82.9)	47 (13.2)	342 (96.1)
	Total	307 (86.2)	49 (13.8)	356 (100)

$N = 44$		Column player		Total
(11.0)		Y	P	
Row player	Y	10 (22.7)	3 (6.8)	13 (29.5)
	P	16 (36.4)	15 (34.1)	31 (70.5)
	Total	26 (59.1)	18 (40.9)	44 (100)

Table 11
Treatment Control II: Absolute and relative (in parentheses) frequencies of outcomes in the last 16 periods

$N = 227$		Column player		Total
(56.8)		Y	P	
Row player	Y	10 (4.4)	7 (3.1)	17 (7.5)
	P	133 (58.6)	77 (33.9)	210 (92.5)
	Total	143 (63.0)	84 (37.0)	227 (100)

$N = 173$		Column player		Total
(43.2)		Y	P	
Row player	Y	55 (31.8)	92 (53.2)	147 (85.0)
	P	9 (5.2)	17 (9.8)	26 (15.0)
	Total	64 (37.0)	109 (63.0)	173 (100)

[23] The increase is significant at $p = 0.004$ (one-tailed MWU test). Note that row players in the BoS subgame play P in 96.1% of all cases (compared to 65% in treatment BoS). Column players' relative frequency of Y rises from 35% (in treatment BoS) to 86.3%. As one-tailed MWU tests indicate, the changes in both row and column players' behavior is significant at $p = 0.004$.

57% of all cases.[24] However, despite the first choice being materially irrelevant, we observe that the incidence of the row player's preferred equilibrium rises from 22.75% in treatment BoS to 58.6% in the left subgame and to 53.2% in the right subgame! Both increases are statistically significant.[25] Comparing the second control treatment with Burn we find that first movers did slightly better in the latter where they got their preferred equilibrium in 68.8% cases conditional on reaching the BoS subgame. However, MWU tests reveal that the differences between the occurrence of the row player's preferred outcome conditional on reaching the BoS subgame in treatment Burn and conditional on reaching the two subgames of the second control treatment are statistically insignificant.[26]

We, therefore, conclude with

Result 3. The data reject both the Forward Induction and the Behavioral Forward Induction Hypothesis but lend significant support to the Physical Timing Hypothesis.

5. Toward a unified theory of physical time in games

The data reported here present strong evidence for the relevance of physical time in games. In some sense the evidence is even stronger than previous evidence from the analysis of the so-called positional order protocol. There, letting players move sequentially in physical time without information revelation induces a strong first-mover advantage as first shown by Cooper et al. (1993). Theoretical attempts to make sense out of such results go back to Amershi et al.'s (1989a, 1989b, 1989c) "manipulated Nash equilibrium". In essence, Amershi et al. refine Nash equilibrium by introducing the notion of "virtual observability": If players move sequentially in physical time, they analyze the extensive form of a game with the same game tree but perfect information. If this new game has a subgame perfect equilibrium outcome that is preferred by the first mover and that coincides with an equilibrium outcome of the original game, players play according to it. If not, physical timing is irrelevant (see also Weber et al., 2004).[27]

However, in our second control treatment "virtual observability" does *not* predict a first mover advantage. After all, players do move simultaneously in both subgames. Hence, our results show that the effect of physical timing goes beyond that of "virtual observability". Rather it seems that players use any clue that is provided by the physical sequence of the play in order to select among multiple equilibria. Hence, even (irrelevant) decisions in the past can render one player's preferred outcome "focal". The data also show which player's: It is always the first mover who has an advantage.

All of this is in line with simple yet widespread social norms like first-come-first-serve. Why such norms have arisen is, of course, an issue we cannot address with our data.

[24] This slight tendency to prefer the left subgame over the right subgame is probably due to some cultural norm and should probably not be a source of concern.

[25] A one-tailed MWU test delivers $p = 0.004$ for the left subgame and $p = 0.0755$ for the right subgame.

[26] A two-tailed MWU tests delivers $p = 0.421$ for the left subgame and $p = 0.310$ for the right subgame.

[27] Note that "virtual observability" selects the first mover's preferred equilibrium in the above-mentioned positional-order-protocol experiment by Cooper et al. (1993).

Nevertheless, we can speculate a little. First-come-first-serve is a simple yet often efficient rule. In particular, it is much more attractive than the opposite rule which prefers late comers. So it might not be that surprising that first-come-first-serve has evolved as a social norm or is used as an explicit allocation mechanism (e.g., for concert ticket sales). And once it has evolved it might be quite natural to "use" it for new problems even if they are of a slightly different nature. In particular, players in the above games might reason that there is a solution to *their* conflict that is often used to resolve other, somewhat similar conflicts: Whoever is first, gets what he wants. A theoretical model of this argument would require formal notions of similarities between problems and rules. Equipped with such notions one could study how a social norm that has evolved for a particular class of problems can cause "spillovers" to other classes. With such spillovers the role of time in one game may crucially depend on its role in another, different game. The evidence we have seen so far suggests that this might indeed be the case.

6. Summary and discussion

In this paper we report the results of an experiment designed to assess the behavioral impact of giving one player the opportunity to burn money prior to the play of a battle-of-the-sexes game. By using forward induction arguments or, alternatively, deletion of iteratively dominated strategies, theory predicts that this variation of the BoS game completely solves the coordination problem. In fact, we find that the option to burn money changes play in the BoS subgame; yet not quite to the extent theory predicts. Furthermore, as in several other studies,[28] we find dramatic differences concerning the play in the extensive and the normal form representation suggesting that having one of the players moving before the other creates a "pseudo" first-mover advantage. This explanation found strong support in two control treatments where a game theoretically irrelevant choice between two subgames still induced a massive advantage for the player who had to make this first choice. The results of the control treatments are related to previous experimental data showing that physical timing can be of utter importance even if game theory predicts it to be irrelevant. In our view, this suggests that further theoretical work to incorporate physical time into game theory could become very fruitful indeed.

Acknowledgments

We are indebted to Sophie Bade, Heike Harmgart, Georg Weizsäcker, an associate editor, three anonymous referees, and seminar participants at NYU and the GSB at the University of Chicago for many helpful comments. Special thanks to Brian Wallace for helping to conduct the experiments. The first author acknowledges financial support from the Economic and Social Research Council (UK) via ELSE. The second author

[28] See e.g. Schotter et al. (1994), Cooper et al. (1993) or Güth et al. (2001). See, however, also Brandts and Charness (2000) who report that behavior is more robust in certain classes of games.

acknowledges financial support from the German Science Foundation, DFG. He also thanks the Center for Experimental Social Science (CESS) at New York University for its hospitality.

Appendix A. Instructions

A.1. Instructions of treatment BoS

Welcome to our experiment! Please read these instructions carefully! Do not speak to your neighbors and keep quiet during the entire experiment! In case you have a question raise your hand! We will then come to you.

In this experiment you will repeatedly make decisions. Doing this you can earn money. How much you earn depends on your decisions and on the decisions of other participants. All participants receive the same instructions.

All participants stay anonymous to the experimenter and also to other participants.

In the experimental situation there are two agents. In each round every participant will be randomly matched with another one. Both participants have to decide simultaneously what to do. In particular, you each choose either “Option 1” or “Option 2”.

Table A.1 is read as follows. The head of a row shows the possible options you have yourself (“Option 1” and “Option 2”). The head of a column shows the identical options of the other participant (also “Option 1” and “Option 2”). For a given combination of decisions, the left number in the box at which row and column intersect is your own payoff and the right number the payoff for the other participant. The payoffs are given in Points. The sum of your individual earnings in the 30 rounds (plus your endowment) determines your monetary payoff in Pounds. After the experiment is finished you will get £1 for every 4 Points that you have earned.

The experiment consists of 30 rounds. In each round you will be randomly matched with one of the other participants. (You do not know with whom you are matched.) Notice that this means that it is not very likely that you will be matched with the same person in two consecutive rounds.

At the start of the experiment you get a one-off endowment of 20 Points. (This is the £5 show-up fee you were promised, see below.)

The sum of your individual earnings in the 30 rounds (plus your endowment) determines your monetary payoff in Pounds. After the experiment is finished you will get £1 for every 4 Points that you have.

A.2. Instructions of treatment Burn/Norm

[In both treatments:] Welcome to our experiment! Please read these instructions carefully! Do not speak to your neighbors and keep quiet during the entire experiment! In case you have a question raise your hand! We will then come to you.

In this experiment you will repeatedly make decisions. Doing this you can earn money. How much you earn depends on your decisions and on the decisions of other participants. All participants receive the same instructions.

All participants stay anonymous to the experimenter and also to other participants.

[Next only for treatment Burn:] In the experimental situation there are two agents called A and B, respectively (see Table A.2). In each round every A will be randomly matched with a B. At the beginning of a round, A has to decide whether he wants to go “to the right” or “to the left”. If he chooses Left then Table L will be relevant; if he chooses Right, Table R will be relevant.

Table A.1

		Other choice	
		Option 1	Option 2
Your	Option 1	0, 0	1, 3
choice	Option 2	3, 1	0, 0

Table A.2

Table L

	Column 1	Column 2
Row 1	0, 0	1, 3
Row 2	3, 1	0, 0

Table R

	Column 1	Column 2
Row 1	−1, 0	0, 3
Row 2	2, 1	−1, 0

Table A.3

		Participant B			
		Column 1	Column 2	Column 3	Column 4
Participant *A*	Row 1	0, 0	0, 0	1, 3	1, 3
	Row 2	3, 1	3, 1	0, 0	0, 0
	Row 3	−1, 0	0, 3	−1, 0	0, 3
	Row 4	2, 1	−1, 0	2, 1	−1, 0

B will be informed about A's decision regarding which table has been chosen. After that A and B have to decide simultaneously about their options in the relevant table. Agent A has to choose either "Row 1" or "Row 2", agent B has to choose either "Column 1" or "Column 2".

The tables are read as follows: The head of a row shows the possible options of agent A ("Row 1" or "Row 2") and the head of a column shows the possible options of agent B ("Column 1" or "Column 2"). For a given combination of decisions, the left number in the box at which row and column intersect is the payoff of A and the right number the one for B. The payoffs are given in Points.

[Next only for treatment Norm:] In the experimental situation there are two agents called A and B, respectively (see Table A.3). In each round every A will be randomly matched with a B. A and B have to decide simultaneously about their options in the table. Agent A has to choose one of the four rows ("Row 1", ..., "Row 4"), agent B has to choose one of the four columns ("Column 1", ..., "Column 4").

The tables are read as follows. The head of a row shows the possible options of agent A ("Row 1" to "Row 4") and the head of a column shows the possible options of agent B ("Column 1" to "Column 4"). For a given combination of decisions, the left number in the box at which row and column intersect is the payoff of A and the right number the one for B. The payoffs are given in Points.

[In both treatments:] The experiment consists of 30 rounds.

For the role assignment and matching procedure the following holds:

(1) In each round you will be matched with another person chosen at random. (You do not know with whom you are matched.)
(2) In each pair of consecutive rounds $(1, 2)$, $(3, 4)$, ... and so on you will be agent A once and agent B once. It will be randomly determined which role you will have in which round. In any case, in the course of the experiment you will be as often in the role of A as you will be in the role of B. (Notice that—because of point 1—it is not very likely that you will be matched with the same other participant during these two rounds.)
(3) In each round an A-participant will always be matched with a B-participant and vice versa. At the start of the experiment you get a one-off endowment of 20 Points. (This is the £5 show-up fee you were promised.)

The sum of your individual earnings in the 30 rounds (plus your endowment) determines your monetary payoff in Pounds. After the experiment is finished you will get £1 for every 4 Points that you have.

References

Amershi, A., Sadanand, A., Sadanand, V., 1989a. Manipulated Nash equilibrium I: Forward induction and thought process dynamics in extensive form. Mimeo.

Amershi, A., Sadanand, A., Sadanand, V., 1989b. Manipulated Nash equilibrium II: Some properties. Mimeo.

Amershi, A., Sadanand, A., Sadanand, V., 1989c. Manipulated Nash equilibrium III: Applications and a preliminary experiment. Mimeo.

Balkenborg, D., 1998. An experiment on forward versus backward induction. Mimeo. University of Southampton.

Ben-Porath, E., Dekel, E., 1992. Signaling future actions and the potential for sacrifice. J. Econ. Theory 57, 36–51.

Brandts, J., Charness, G., 2000. Hot vs. cold: Sequential responses in simple experimental games. Exper. Econ. 2, 227–238.

Brandts, J., Holt, C.A., 1992. Forward induction: Experimental evidence from two-stage games with complete information. In: Isaac, R.M. (Ed.), Research in Experimental Economics, vol. 5. Jai Press, pp. 119–136.

Brandts, J., Holt, C.A., 1995. Limitations of dominance and forward induction: experimental evidence. Econ. Lett. 49, 391–395.

Burton, A., Loomes, G., Sefton, M., 1999. Communication and efficiency in coordination game experiments. CeDEx working paper 1999-1. University of Nottingham.

Cachon, G.P., Camerer, C.F., 1996. Loss-avoidance and forward induction in experimental coordination games. Quart. J. Econ. 111, 165–194.

Charness, G.B., 2000. Self-serving cheap talk and credibility: a test of Aumann's conjecture. Games Econ. Behav. 33, 177–194.

Clark, K., Kay, S., Sefton, M., 2000. When are Nash equilibria self-enforcing: an experimental analysis. Int. J. Game Theory 29, 495–515.

Cooper, R., DeJong, D.V., Forsythe, R., Ross, T.W., 1989. Communication in the battle of the sexes game: some experimental results. RAND J. Econ. 20, 568–587.

Cooper, R., DeJong, D.V., Forsythe, R., Ross, T.W., 1992. Forward induction in coordination games. Econ. Lett. 40, 167–172.

Cooper, R., DeJong, D.V., Forsythe, R., Ross, T.W., 1993. Forward induction in the battle-of-the-sexes game. Amer. Econ. Rev. 83, 1303–1316.

Costa-Gomes, M., Crawford, V., Broseta, B., 2001. Cognition and behavior in normal-form games: an experimental study. Econometrica 69, 1193–1235.

Farrell, J., 1987. Cheap talk, coordination and entry. RAND J. Econ. 18, 34–39.

Fischbacher, U., 1999. Z-Tree, Zurich toolbox for readymade economic experiments. Working paper Nr. 21. Institute for Empirical Research in Economics, University of Zurich.

Güth, W., Huck, S., Müller, W., 2001. The relevance of equal splits in ultimatum games. Games Econ. Behav. 37, 161–169.

Güth, W., Huck, S., Rapoport, A., 1998. The limitations of the positional order effect: can it support silent threats and non-equilibrium behavior? J. Econ. Behav. Organ. 34, 313–325.

Harsanyi, J.C., Selten, R., 1988. A General Theory of Equilibrium Selection in Games. MIT Press, Cambridge, MA.

Kohlberg, E., Mertens, J.-F., 1986. On the strategic stability of equilibria. Econometrica 54, 1003–1038.

Kübler, D., Weizsäcker, G., in press. Limited depth of reasoning and failure of cascade formation in the laboratory. Rev. Econ. Stud.

Muller, R.A., Sadanand, A., 2001. Order of play, forward induction, and presentation effects in two-person games. Mimeo.

Nachbar, J., 1990. Evolutionary selection dynamics in games: convergence and limit properties. Int. J. Game Theory 19, 59–89.

Nagel, R., 1995. Unravelling in guessing games: an experimental study. Amer. Econ. Rev. 85, 1313–1326.

Rapoport, A., 1997. Order of play in strategically equivalent games in extensive form. Int. J. Game Theory 26, 113–136.

Samuelson, L., 1993. Does evolution eliminate dominated strategies? Binmore, K.G., Kirman, A.P., Tanni, P. (Eds.), Frontiers of Game Theory. MIT Press, Cambridge, MA, pp. 213–236.

Schotter, A., Weigelt, K., Wilson, C., 1994. A laboratory investigation of multiperson rationality and presentation effects. Games Econ. Behav. 6, 445–468.

Thompson, F.B., 1952. Equivalence of games in extensive form. Research memorandum RM-759. The RAND Corporation, Santa Monica.

van Damme, E., 1989. Stable equilibria and forward induction. J. Econ. Theory 48, 476–496.

Van Huyck, J., Battalio, R., Beil, R., 1993. Asset markets as an equilibrium selection mechanism: coordination failure, game form auctions, and tacit communication. Games Econ. Behav. 5, 485–504.

von Neumann, J., Morgenstern, O., 1947. Theory of Games and Economic Behavior, 2nd Edition. Princeton Univ. Press, Princeton.

Weber, R., Camerer, C.F., Knez, M., 2004. Timing and virtual observability in ultimatum bargaining and 'weak link' coordination games. Exper. Econ. 7, 25–48.

Weizsäcker, G., 2003. Ignoring the rationality of others: evidence from experimental normal-form games. Games Econ. Behav. 44, 145–171.